Serial Number: 10/540,091

CRF Edit Date: 7/1/05

Edited by: Av

_____ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

_____ Corrected the SEQ ID NO. Sequence numbers edited were:

_____

_____ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

_____

✓ Deleted: ___ invalid beginning/end-of-file text ; ___ page numbers

_____ Inserted mandatory headings/numeric identifiers, specifically:

_____

_____ Moved responses to same line as heading/numeric identifier, specifically:

_____

_____ Other:

_____
_____
_____

Revised 09/09/2003

PCT

```
 3 <110> APPLICANT: Wu, Wenping
 4          Lange, Lene
 5          Skovlund, Dominique Aubert
 6          Liu, Ye
 8 <120> TITLE OF INVENTION: Polypeptides having Cellobiohydrlase II Activity and
 9          Polynucleotides Encoding Same
11 <130> FILE REFERENCE: 10377.204-US
C--> 13 <140> CURRENT APPLICATION NUMBER: US/10/540,091
C--> 13 <141> CURRENT FILING DATE: 2005-06-20
13 <160> NUMBER OF SEQ ID NOS: 36
15 <170> SOFTWARE: PatentIn version 3.3
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 1731
19 <212> TYPE: DNA
20 <213> ORGANISM: Chaetomium thermophilum NP000980
23 <220> FEATURE:
24 <221> NAME/KEY: CDS
25 <222> LOCATION: (63)..(1493)
27 <400> SEQUENCE: 1
28 cggggggggg ggacagcaca acagagtcaa gacaagcttg gtcgctttgt cagaagttca      60
30 tc atg gct aag cag ctg ctg ctc act gcc gct ctt gcg gcc act tcg     107
31    Met Ala Lys Gln Leu Leu Leu Thr Ala Ala Leu Ala Ala Thr Ser
32    1               5                  10                  15
34 ctg gct gcc cct ctc ctt gag gag cgc cag agc tgc tcc tcc gtc tgg     155
35 Leu Ala Ala Pro Leu Leu Glu Glu Arg Gln Ser Cys Ser Ser Val Trp
36                  20                  25                  30
38 ggt caa tgc ggt ggc atc aat tac aac ggc ccg acc tgc tgc cag tcc     203
39 Gly Gln Cys Gly Gly Ile Asn Tyr Asn Gly Pro Thr Cys Cys Gln Ser
40                  35                  40                  45
42 ggc agt gtt tgc act tac ctg aat gac tgg tac agc cag tgc att ccc     251
43 Gly Ser Val Cys Thr Tyr Leu Asn Asp Trp Tyr Ser Gln Cys Ile Pro
44          50                  55                  60
46 ggt cag gct cag ccc ggc acg act agc acc acg gct cgg acc acc agc     299
47 Gly Gln Ala Gln Pro Gly Thr Thr Ser Thr Thr Ala Arg Thr Thr Ser
48    65                  70                  75
50 acc agc acc acc agc act tcg tcg gtc cgc ccg acc acc tcg aat acc     347
51 Thr Ser Thr Thr Ser Thr Ser Ser Val Arg Pro Thr Thr Ser Asn Thr
52 80                  85                  90                  95
54 cct gtg acg act gct ccc ccg acg acc acc atc ccg ggc ggc gcc tcg     395
55 Pro Val Thr Thr Ala Pro Pro Thr Thr Thr Ile Pro Gly Gly Ala Ser
56                  100                 105                 110
58 agc acg gcc agc tac aac ggc aac ccg ttt tcg ggt gtt caa ctt tgg     443
59 Ser Thr Ala Ser Tyr Asn Gly Asn Pro Phe Ser Gly Val Gln Leu Trp
```

```
60              115               120               125
62 gcc aac acc tac tac tcg tcc gag gtg cac act ttg gcc atc ccc agc      491
63 Ala Asn Thr Tyr Tyr Ser Ser Glu Val His Thr Leu Ala Ile Pro Ser
64         130               135               140
66 ttg tct cct gag ctg gct gcc aag gcc gcc aag gtc gct gag gtt ccc      539
67 Leu Ser Pro Glu Leu Ala Ala Lys Ala Ala Lys Val Ala Glu Val Pro
68      145               150               155
70 agc ttc cag tgg ctc gac cgc aat gtg act gtt gac act ctc ttc tcc      587
71 Ser Phe Gln Trp Leu Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser
72 160               165               170               175
74 ggc act ctt gcc gaa atc cgc gcc gcc aac cag cgc ggt gcc aac ccg      635
75 Gly Thr Leu Ala Glu Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro
76               180               185               190
78 cct tat gcc ggc att ttc gtg gtt tat gac tta cca gac cgt gat tgc      683
79 Pro Tyr Ala Gly Ile Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys
80               195               200               205
82 gcg gct gct gct tcg aac ggc gag tgg tct atc gcc aac aat ggt gcc      731
83 Ala Ala Ala Ala Ser Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala
84         210               215               220
86 aac aac tac aag cgc tac atc gac cgg atc cgt gag ctc ctt atc cag      779
87 Asn Asn Tyr Lys Arg Tyr Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln
88      225               230               235
90 tac tcc gat atc cgc act att ctg gtc att gaa cct gat tcc ctg gcc      827
91 Tyr Ser Asp Ile Arg Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala
92 240               245               250               255
94 aac atg gtc acc aac atg aac gtc cag aag tgc tcg aac gct gcc tcc      875
95 Asn Met Val Thr Asn Met Asn Val Gln Lys Cys Ser Asn Ala Ala Ser
96               260               265               270
98 act tac aag gag ctt act gtc tat gcc ctc aaa cag ctc aat ctt cct      923
99 Thr Tyr Lys Glu Leu Thr Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro
100            275               280               285
102 cac gtt gcc atg tac atg gat gct ggc cac gct ggc tgg ctt ggc tgg      971
103 His Val Ala Met Tyr Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp
104            290               295               300
106 ccc gcc aac atc cag cct gct gct gag ctc ttt gct caa atc tac cgc     1019
107 Pro Ala Asn Ile Gln Pro Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg
108      305               310               315
110 gac gct ggc agg ccc gct gct gtc cgc ggt ctt gcg acc aac gtt gcc     1067
111 Asp Ala Gly Arg Pro Ala Ala Val Arg Gly Leu Ala Thr Asn Val Ala
112 320               325               330               335
114 aac tac aat gct tgg tcg atc gcc agc cct ccg tcc tac acc tct cct     1115
115 Asn Tyr Asn Ala Trp Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro
116               340               345               350
118 aac ccg aac tac gac gag aag cac tat att gag gcc ttt gct cct ctt     1163
119 Asn Pro Asn Tyr Asp Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu
120            355               360               365
122 ctc cgc aac cag ggc ttc gac gca aag ttc atc gtc gac acc ggc cgt     1211
123 Leu Arg Asn Gln Gly Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg
124            370               375               380
```

Input Set : **A:\PTO.AMC.txt**
Output Set: **N:\CRF4\07012005\J540091.raw**

```
126 aac ggc aag cag ccc act ggc cag ctt gaa tgg ggt cac tgg tgc aat    1259
127 Asn Gly Lys Gln Pro Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn
128     385                 390                 395
130 gtc aag gga act ggc ttc ggt gtg cgc cct act gct aac act ggg cat    1307
131 Val Lys Gly Thr Gly Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His
132 400                 405   .             410                 415
134 gaa ctt gtt gat gct ttc gtg tgg gtc aag ccc ggt ggc gag tcc gac    1355
135 Glu Leu Val Asp Ala Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp
136                 420                 425                 430
138 ggc acc agt gcg gac acc agc gct gct cgt tat gac tat cac tgc ggc    1403
139 Gly Thr Ser Ala Asp Thr Ser Ala Ala Arg Tyr Asp Tyr His Cys Gly
140                 435                 440                 445
142 ctt tcc gac gca ctg act ccg gcg cct gag gct ggc caa tgg ttc cag    1451
143 Leu Ser Asp Ala Leu Thr Pro Ala Pro Glu Ala Gly Gln Trp Phe Gln
144                 450                 455                 460
146 gct tat ttc gaa cag ctg ctc atc aat gcc aac cct ccg ctc            1493
147 Ala Tyr Phe Glu Gln Leu Leu Ile Asn Ala Asn Pro Pro Leu
148     465                 470                 475
150 tgaacggaag cggagatacc ggaaggcggt gagaagagcg gaattcaagt ctgcttatca  1553
152 aaatccactc accaagtgga ttaaagcgga tttatacatc tgagaaacaa cctgctttaa  1613
154 actcttcttg tacatatttc acttcgagac gtgcctcttt ctcaggagca ctgtagatac  1673
156 caatatatct gtcacatttc atataaaaaa aaaaaaaaag aaaaaagta ctagtcga     1731
159 <210> SEQ ID NO: 2
160 <211> LENGTH: 477
161 <212> TYPE: PRT
162 <213> ORGANISM: Chaetomium thermophilum NP000980
164 <400> SEQUENCE: 2
166 Met Ala Lys Gln Leu Leu Leu Thr Ala Ala Leu Ala Ala Thr Ser Leu
167 1                 5                 10                 15
170 Ala Ala Pro Leu Leu Glu Glu Arg Gln Ser Cys Ser Ser Val Trp Gly
171                 20                 25                 30
174 Gln Cys Gly Gly Ile Asn Tyr Asn Gly Pro Thr Cys Cys Gln Ser Gly
175             35                 40                 45
178 Ser Val Cys Thr Tyr Leu Asn Asp Trp Tyr Ser Gln Cys Ile Pro Gly
179     50                 55                 60
182 Gln Ala Gln Pro Gly Thr Thr Ser Thr Thr Ala Arg Thr Thr Ser Thr
183 65                 70                 75                 80
186 Ser Thr Thr Ser Thr Ser Ser Val Arg Pro Thr Thr Ser Asn Thr Pro
187                 85                 90                 95
190 Val Thr Thr Ala Pro Pro Thr Thr Thr Ile Pro Gly Gly Ala Ser Ser
191                 100                 105                 110
194 Thr Ala Ser Tyr Asn Gly Asn Pro Phe Ser Gly Val Gln Leu Trp Ala
195             115                 120                 125
198 Asn Thr Tyr Tyr Ser Ser Glu Val His Thr Leu Ala Ile Pro Ser Leu
199     130                 135                 140
202 Ser Pro Glu Leu Ala Ala Lys Ala Ala Lys Val Ala Glu Val Pro Ser
203 145                 150                 155                 160
206 Phe Gln Trp Leu Asp Arg Asn Val Thr Val Asp Thr Leu Phe Ser Gly
207                 165                 170                 175
```

```
210 Thr Leu Ala Glu Ile Arg Ala Ala Asn Gln Arg Gly Ala Asn Pro Pro
211             180             185             190
214 Tyr Ala Gly Ile Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys Ala
215         195             200             205
218 Ala Ala Ala Ser Asn Gly Glu Trp Ser Ile Ala Asn Asn Gly Ala Asn
219     210             215             220
222 Asn Tyr Lys Arg Tyr Ile Asp Arg Ile Arg Glu Leu Leu Ile Gln Tyr
223 225             230             235             240
226 Ser Asp Ile Arg Thr Ile Leu Val Ile Glu Pro Asp Ser Leu Ala Asn
227         245             250             255
230 Met Val Thr Asn Met Asn Val Gln Lys Cys Ser Asn Ala Ala Ser Thr
231         260             265             270
234 Tyr Lys Glu Leu Thr Val Tyr Ala Leu Lys Gln Leu Asn Leu Pro His
235         275             280             285
238 Val Ala Met Tyr Met Asp Ala Gly His Ala Gly Trp Leu Gly Trp Pro
239     290             295             300
242 Ala Asn Ile Gln Pro Ala Ala Glu Leu Phe Ala Gln Ile Tyr Arg Asp
243 305             310             315             320
246 Ala Gly Arg Pro Ala Ala Val Arg Gly Leu Ala Thr Asn Val Ala Asn
247         325             330             335
250 Tyr Asn Ala Trp Ser Ile Ala Ser Pro Pro Ser Tyr Thr Ser Pro Asn
251         340             345             350
254 Pro Asn Tyr Asp Glu Lys His Tyr Ile Glu Ala Phe Ala Pro Leu Leu
255         355             360             365
258 Arg Asn Gln Gly Phe Asp Ala Lys Phe Ile Val Asp Thr Gly Arg Asn
259     370             375             380
262 Gly Lys Gln Pro Thr Gly Gln Leu Glu Trp Gly His Trp Cys Asn Val
263 385             390             395             400
266 Lys Gly Thr Gly Phe Gly Val Arg Pro Thr Ala Asn Thr Gly His Glu
267             405             410             415
270 Leu Val Asp Ala Phe Val Trp Val Lys Pro Gly Gly Glu Ser Asp Gly
271         420             425             430
274 Thr Ser Ala Asp Thr Ser Ala Ala Arg Tyr Asp Tyr His Cys Gly Leu
275         435             440             445
278 Ser Asp Ala Leu Thr Pro Ala Pro Glu Ala Gly Gln Trp Phe Gln Ala
279     450             455             460
282 Tyr Phe Glu Gln Leu Leu Ile Asn Ala Asn Pro Pro Leu
283 465             470             475
286 <210> SEQ ID NO: 3
287 <211> LENGTH: 1272
288 <212> TYPE: DNA
289 <213> ORGANISM: Myceliophthora thermophila
292 <220> FEATURE:
293 <221> NAME/KEY: misc_feature
294 <222> LOCATION: (61)..(61)
295 <223> OTHER INFORMATION: n is a, c, g, or t
297 <220> FEATURE:
298 <221> NAME/KEY: misc_feature
299 <222> LOCATION: (87)..(87)
```

RAW SEQUENCE LISTING   DATE: 07/01/2005
PATENT APPLICATION: US/10/540,091   TIME: 12:43:14

Input Set : A:\PTO.AMC.txt
Output Set: N:\CRF4\07012005\J540091.raw

```
      300 <223> OTHER INFORMATION: n is a, c, g, or t
      302 <220> FEATURE:
      303 <221> NAME/KEY: misc_feature
      304 <222> LOCATION: (420)..(420)
      305 <223> OTHER INFORMATION: n is a, c, g, or t
      307 <400> SEQUENCE: 3
      308 acgtggatcc gaattcaagc ttcatggggt tctcctgcga tccactagta acggctcgcc    60
W-->  310 ngagagctgg aaagccagca gcacctnttg gagaggcagc tctctctctc caccacgccc   120
      312 ccgcccgtct ccagccctcg tgaccagcat tcccggcggt gcgacctcca cggcgagcta   180
      314 ctctggcaac cccttctcgg gcgtccggct cttcgccaac gactactaca ggtccgaggt   240
      316 ccacaatctc gccattccta gcatgactgg tactctggcg gctcaaggct tccgccgtcg   300
      318 cgcgaagtcc ctagcttcca gtggctcgac acggaacgtg cactcatcag acaccctgat   360
      320 ggtccagact ctgtacccag gtccgggctc tcaataaggc acggtgaaca atcctacccn   420
      322 tatgctgccc aactcgtcgt ctacgacctc cccgaccgtg actgtccgc cgctgcgtcc   480
      324 aacggygagt tttcgattgc aaacggcggc gccgccaact acaggagcta catcgacgct   540
      326 atccgcaagc acatcattga gtactcggac atccggatca tcctggttat cgagcccgac   600
      328 tcgatggcca acatggtgac caacatgaac gtggccaagt gcagcaacgc cgcgtcgacg   660
      330 taccacgagt tgaccgtgta cgcgctcaag cagctgaacc tgcccaacgt cgccatgtat   720
      332 ctcgacgccg gccacgccgg ctggctcggc tggcccgcca acatccagcc cgccgccgag   780
      334 ctgtttgccg gcatctacaa tgatgccggc aagccggctg ccgtccgcgg cctggccact   840
      336 aacgtcgcca actacaacgc ctggagcatc gcttcggccc cgtcgtacac gtcggctaac   900
      338 cctaactacg acgagaagca ctacatcgag gccttcagcc cgctcttgaa ctcggccggc   960
      340 ttccccgcac gcttcattgt cgacactggc cgcaacggca acaacctac cggccaacaa  1020
      342 cagtggggcg actggtgcaa tgtcaagggc accggctttg gcgtgcgccc gacggccaac  1080
      344 acgggccacg agctggtcga tgcctttgtc tgggtcaagc ccggcggcga gtccgacggc  1140
      346 acaagcgaca ccagcgccgc ccgctacgac taccactgcg gcctgtccga tgccctgcag  1200
      348 cctgcccccg aggctggaca gtggttccag gcctacttcg agcagctgct caccaacgcc  1260
      350 aacccgcccct tc                                                     1272
      353 <210> SEQ ID NO: 4
      354 <211> LENGTH: 420
      355 <212> TYPE: PRT
      356 <213> ORGANISM: Myceliophthora thermophila
      358 <400> SEQUENCE: 4
      360 Thr Trp Ile Arg Ile Gln Ala Ser Trp Gly Ser Pro Ala Ile His Arg
      361 1               5                  10                  15
      364 Leu Ala Glu Leu Glu Ser Gln Gln His Leu Leu Glu Arg Gln Leu Ser
      365             20                  25                  30
      368 Leu Ser Thr Thr Pro Pro Pro Val Ser Ser Pro Arg Asp Gln His Ser
      369             35                  40                  45
      372 Arg Arg Cys Asp Leu His Gly Glu Leu Leu Trp Gln Pro Leu Leu Gly
      373         50                  55                  60
      376 Arg Pro Ala Leu Arg Gln Arg Leu Leu Gln Val Arg Gly Pro Gln Ser
      377 65                  70                  75                  80
      380 Arg His Ser His Asp Trp Tyr Ser Gly Gly Ser Arg Leu Pro Pro Ser
      381                 85                  90                  95
      384 Arg Glu Val Pro Ser Phe Gln Trp Leu Asp Thr Glu Arg Ala Leu Ile
      385             100                 105                 110
      388 Arg His Pro Asp Gly Pro Asp Ser Val Pro Arg Ser Gly Leu Ser Ile
      389             115                 120                 125
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 07/01/2005
PATENT APPLICATION: US/10/540,091   TIME: 12:43:15

Input Set : A:\PTO.AMC.txt
Output Set: N:\CRF4\07012005\J540091.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 61,87,420
Seq#:15; N Pos. 903,1011,1017
Seq#:15; Xaa Pos. 217,264,280,301,339,355
Seq#:16; Xaa Pos. 217,264,280,301,339,355
Seq#:19; N Pos. 36
Seq#:20; Xaa Pos. 20,28,31,35,42,48,60
Seq#:27; N Pos. 6,15
Seq#:28; N Pos. 6,9,15
Seq#:29; N Pos. 3,9,12
Seq#:30; N Pos. 12
Seq#:31; N Pos. 6,12
Seq#:32; N Pos. 6

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:27,28,29,30,31,32,33,34,35,36

VERIFICATION SUMMARY
PATENT APPLICATION: **US/10/540,091**

DATE: 07/01/2005
TIME: 12:43:15

Input Set : **A:\PTO.AMC.txt**
Output Set: **N:\CRF4\07012005\J540091.raw**

```
L:13 M:270 C: Current Application Number differs, Replaced Current Application No
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:310 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:60
M:341 Repeated in SeqNo=3
L:970 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:672
M:341 Repeated in SeqNo=15
L:1111 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:208
M:341 Repeated in SeqNo=16
L:1266 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:0
L:1292 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:16
M:341 Repeated in SeqNo=20
L:1956 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:0
L:1976 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:0
L:1996 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:0
L:2016 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:0
L:2038 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31 after pos.:0
L:2060 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32 after pos.:0
```